ROBUST LOUDSPEAKER EQUALIZATION BASED ON POSITION-INDEPENDENT EXCESS PHASE MODELING

Lars-Johan Brännmark and Anders Ahlén

Uppsala University
Dept. of Engineering Sciences, Signals and Systems Group
PO Box 534, SE-751 21 Uppsala, Sweden

ABSTRACT

A well known problem in loudspeaker equalization is that mixed phase design of the inverse filter causes residual "pre-ringings" in the equalized system, due to the spatial variability of loudspeaker-room transfer functions. A common strategy for robust and perceptually acceptable equalization is therefore to use minimum phase filters only. In this paper, a method for cautious mixed phase equalization is proposed. By analysis of a set of room transfer functions, it is concluded that some non-minimum phase zeros are insensitive to receiver position, and can therefore be robustly inverted. The method improves upon a minimum phase equalization by extending the minimum phase model with a robustly invertible all pass link. Validation measurements show that the time-domain aspect of equalization is improved throughout the spatial region of interest, while pre-ringings are kept at a very low and prespecified level.

Index Terms— Loudspeaker equalization, Robustness, All-pass, Excess phase, Acoustic signal processing

1. INTRODUCTION

The aim of single-channel loudspeaker equalization is to construct a linear time-invariant filter which improves sound reproduction over a spatial region, based on a set of Room Transfer Function (RTF) models acquired within the region. This is known to be a hard problem due to the spatial variability of the RTFs. In particular, if a filter is designed to compensate for non minimum phase behavior—so called *mixed phase* equalization—it will necessarily contain a non-causal pre-response which may not exactly match any true RTF in the room. As a result, the equalized system will contain residual pre-ringings.

Several different approaches for robust equalization have been suggested in the literature. Examples include multiple-point least-squares (LS) methods [1, 2], common acoustical poles [3], or complex smoothing [4]. LS and complex smoothing approaches are mixed phase methods, which however lack explicit control of pre-ringings. Most other methods use minimum phase modeling and subsequent inversion, which is free from pre-ringing artifacts but lacks precision in terms of phase response correction.

The purpose of this paper is to present a new method for cautious mixed phase equalization, yielding spatially robust inversion without the adverse pre-ringing side effects. This is accomplished by concatenating a minimum phase inverse with a robustly designed all pass filter. The minimum phase inverse used here is a typical "magnitude correction" filter of moderate spectral resolution.

First author is also with Dirac Research AB, Hansellisgatan 6, SE-754 50 Uppsala, Sweden

The paper is organized as follows. In Section 2, we discuss how RTF zeros vary with receiver position, providing a rationale for our mixed phase inversion approach. In Section 3, a quantitative relation is established between the variability of non-minimum phase zeros on the one hand, and residual pre-echoes on the other. As an outcome of this analysis, a strategy for robust mixed phase equalization is proposed. In Section 4, the performance of the new equalizer is evaluated using real measurements in the listener region. Section 5 concludes the paper and points out some directions for further improvement.

2. VARIABILITY OF RTF ZEROS

In order to obtain a qualitative understanding of how RTF zeros vary with receiver position, a set of 18 RTFs were acquired along a line of microphone positions, spaced 2.5 cm apart. The reverberation time in the room was about 0.4 sec in the low frequency region, motivating an FIR model covering this time frame. As is evident from

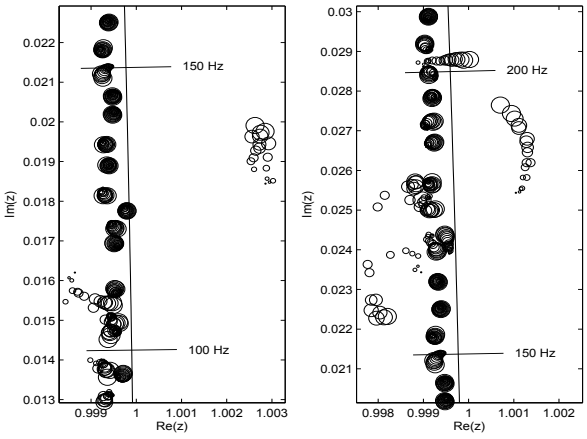

Fig. 1. Zeros of 18 transfer functions, acquired from microphone positions along a line. Zeros are represented by circles, where different radii are used to distinguish individual microphone positions. The smallest and largest radii represent the start and stop positions, respectively. The two diagrams represent zoomed segments of the complex plane near the unit circle, at $100 - 150$ Hz (left) and $150 - 200$ Hz (right).

Figure 1, zeros of the RTFs move around as the microphone position changes. In particular, slightly above 200 Hz, a zero moves from the

inside of the unit circle to the outside—a typical example of a zero which cannot be inverted without causing severe pre- or post-ringing errors in most positions. However, some zeros further out from the unit circle exhibit a more static behaviour. For example, in the left diagram of Figure 1 there is a moving zero outside the unit circle at about 135 Hz. This zero is more static and further away from the unit circle than the zero present at about 185 Hz in the right diagram. Since the former causes less variation along the unit circle than the latter, inversion of the former results in less error than inversion of the latter. Based on these observations, it is reasonable to assume that some non-minimum phase zeros can be safely inverted under a constraint of a maximum tolerable residual pre-ringing.

3. MINIMUM PHASE/ALL-PASS MODELING AND INVERSION

Any linear time-invariant system with transfer function $H(z)$ can be expressed as a cascade of a minimum phase filter and an all pass filter, $H(z) = H_m(z)H_{ap}(z)$. The use of such a decomposition simplifies the analysis and it separates the minimum and excess phase parts of the inverse filter design. Clearly the magnitude is completely contained in the minimum phase part and the all pass part is determined from the non-minimum phase zeros of $H(z)$ [5, 6]. The causal and stable inverse $G_m(z)$ of the minimum phase part $H_m(z)$ is obtained by simply inverting its transfer function: $G_m(z) = 1/H_m(z)$. Similarly, the exact inverse of $H_{ap}(z)$ is the noncausal all pass filter obtained as $G_{ap}(z) = 1/H_{ap}(z)$. In the time domain, this is equivalent to time-reversing the all pass impulse response, $g_{ap}(k) = h_{ap}(-k)$. $G_{ap}(z)$ constructed this way is, of course, not realizable, but it can be approximated with arbitrary accuracy by truncating its impulse response after time-shifting it by as many samples as necessary. The total inverse filter will thus contain the time-reversed tail of an all pass filter. Applied to a set of RTFs, such a filter will cause residual pre-ringings, should the zeros of the RTFs differ significantly from the model on which the filter was based. This effect will be analyzed next.

3.1. Analysis of pre-ringing error

Suppose that a noncausal filter with transfer function $G(z)$ has been designed to be the inverse of a system $H(z)$, but with a small mismatch, so that the poles of $G(z)$ do not completely cancel the zeros of $H(z)$. The residual pre-ringing that results can be quantified as follows.

Let a zero of $H(z)$ be represented by $z_0 = r_0 e^{i\omega_0}$ and a perturbation to this zero by $\epsilon = \rho e^{i\theta}$ where $r_0 > 1$; $0 < \rho \ll 1$; $0 < \omega_0 < \pi$; $-\pi \leq \theta \leq \pi$. Suppose that $H(z)$ contains a complex conjugate pair of zeros at $z_0 + \epsilon$ and $\overline{z_0 + \epsilon}$, so that

$$H(z) = H_1(z)H_2(z) = (z - (z_0 + \epsilon))(z - (\overline{z_0 + \epsilon}))H_2(z)\,. \quad (1)$$

Furthermore, suppose that $G(z)$ contains the pole pair z_0 and $\overline{z_0}$,

$$G(z) = G_1(z)G_2(z) = \frac{1}{(z - z_0)(z - \overline{z_0})}G_2(z). \quad (2)$$

The total transfer function of the equalized system thus becomes

$$\begin{aligned} H_{tot}(z) &= \frac{(z - (z_0 + \epsilon))(z - (\overline{z_0 + \epsilon}))}{(z - z_0)(z - \overline{z_0})}G_2(z)H_2(z) \\ &= \left(1 - \frac{(\epsilon + \bar{\epsilon})z - (\epsilon\overline{z_0} + \bar{\epsilon}z_0 + \epsilon\bar{\epsilon})}{(z - z_0)(z - \overline{z_0})}\right)G_2(z)H_2(z) \\ &= \left(1 - 2\Re(\epsilon)\frac{z - \frac{|z_0+\epsilon|^2 - |z_0|^2}{2\Re(\epsilon)}}{(z - z_0)(z - \overline{z_0})}\right)G_2(z)H_2(z). \quad (3) \end{aligned}$$

Applying the inverse z-transform on each factor in the last line of (3) yields

$$h_{tot}(k) = \Big[\delta(k) + Ar_0^k \cos(-\omega_0 k + \Phi)u(-k)\Big] * g_2(k) * h_2(k), \quad (4)$$

where $\delta(k)$ is the Kronecker delta function, $u(k)$ is the unit step function, and

$$\Phi = \arctan\left(\frac{\frac{2\Re(\epsilon)|z_0|^2}{|z_0+\epsilon|^2 - |z_0|^2} - \Re(z_0)}{\Im(z_0)}\right), \quad (5)$$

$$A = \frac{|z_0 + \epsilon|^2 - |z_0|^2}{|z_0|^2 \cos\Phi}. \quad (6)$$

In the above, $\Re(z_0)$ and $\Im(z_0)$ denote real and imaginary parts of z_0. In (3) and (5) we have used the assumption that $\Re(\epsilon) \neq 0$, and $|z_0 + \epsilon|^2 \neq |z_0|^2$, which is reasonable for measured data. Equation (4) clearly shows how the pole/zero mismatch ϵ between $H(z)$ and $G(z)$ has created a noncausal ringing which affects the total system in a convolutive way.

Suppose now that $H_n(z)$; $n \in \{1, ..., N\}$ represent a set of N RTFs, each containing M zeros z_{nm}; $m \in \{1, ..., M\}$. Furthermore suppose that these zeros are expressed as perturbations, $z_{nm} = z_{0m} + \epsilon_{nm}$, of the *nominal zeros* z_{0m}; $m \in \{1, ..., M_o, M_o + 1, ..., M\}$, where the first M_o nominal zeros are located outside the unit circle in the upper half plane. Once the nominal zeros $z_{0m} = r_{0m}e^{i\omega_{0m}}$ and their perturbations $\epsilon_{nm} = \rho_{nm}e^{i\theta_{nm}}$ have been determined, equations (4), (5) and (6) with obvious modifications can be used to determine the maximum amplitudes $A_1, ..., A_{M_o}$ of the residual pre-ringings caused by placing poles at the nominal zero locations $z_{01}, ..., z_{0M_o}$ and their conjugated counterparts $\overline{z}_{01}, ..., \overline{z}_{0M_o}$.

3.2. Extraction of position-independent excess phase zeros

The next step towards obtaining a robust mixed phase inverse is to actually determine the nominal zeros of interest and their perturbations. If the RTF zeros are located close to one another in clusters outside the unit circle, they are regarded as position-independent, and can be safely inverted. A procedure for finding such clusters is outlined below.

1. From the zeros of N RTFs, find M_o well separated clusters of non-minimum phase zeros (if such clusters exist) in the upper half plane, where each cluster contains exactly one zero from each RTF.
2. Let e.g. the arithmetic mean of the zeros in the m^{th} cluster represent a nominal zero location z_{0m}; $m \in \{1, ..., M_o\}$.
3. For each zero z_{nm} in cluster m, construct the perturbation $\epsilon_{nm} = z_{nm} - z_{0m}$; $n \in \{1, ..., N\}$, and use equations (4), (5) and (6) to compute the maximum pre-ringing amplitude A_m.

4. Decide if the zeros in cluster $m \in \{1, ..., M_o\}$ are invertible by a pole at z_{0m}, using a pre-ringing envelope constraint. A reasonable constraint would, for example, be that for any response in the listening region, the pre-ringing level should be at least 60 dB below the maximum peak level for all time instants before -5 msec, with the maximum peak level at 0 msec. At a sampling rate of 44100 Hz, this means that $20\log_{10}(A_m r_{0m}^{-220}) < -60$. If this constraint is fulfilled, accept the nominal zero z_{0m}, referring to cluster m, as a position-independent, robustly invertible zero. Otherwise, reject the cluster.

Note that the above procedure merely illustrates the steps required to extract robustly invertible zeros.

3.3. Construction of inverse filters

With $H(z)$ parameterized as $H_m(z)H_{ap}(z)$ we will use a two-step procedure to obtain a robust inverse. First, a conventional way to obtain a robust minimum phase inverse is used. Thus a minimum phase model $H_m(z)$ is constructed from 9 RTFs, using power response averaging, magnitude regularization below 40 Hz and above 20 kHz, and $1/6$ octave smoothing of the resulting magnitude curve. The minimum phase inverse is then obtained as $1/H_m(z)$ For details about this kind of inverse filter construction, see e.g. [6, 7, 8]. Second, having obtained a set of position-insensitive non-minimum phase zeros defining the all pass link $H_{ap}(z)$, common to all RTFs in the listener region, inversion of this link is expected not only to improve phase response in the region, but also to correct for phase errors possibly introduced by the minimum phase inverse. The robust all pass inverse is obtained by truncating and time-reversing the impulse response of $H_{ap}(z)$. The total mixed phase inverse is obtained by convolving the minimum phase and all pass inverses.

4. EQUALIZATION PERFORMANCE

4.1. Experimental conditions

In a room of dimensions $4.5 \times 6 \times 2.6$ m, with a distance between loudspeaker and microphones equal to 2.5 m, 9 measurement positions for filter design, and 9 for validation were selected according to Figure 2. This microphone configuration is designed to cover the typical head movements of a normal listener. The measurement points were separated into a design set and a validation set, to ensure that the equalizer is not over-fitted to a particular set of responses, causing improvement at these particular points only. A test signal was prefiltered with the minimum phase and mixed phase equalizer filters designed as in subsection 3.3, and validation impulse responses of length 0.5 sec were acquired in the black microphone positions of Figure 2. In the mixed phase design, the pre-ringing constraint described in subsection 3.2 was used.

4.2. Methods for evaluation

In order to correctly assess the robustness of equalization, it is necessary not to focus on the behaviour in a single point, but on the average (or sometimes, worst case) behaviour in the whole set of validation points. We shall use the average Schroeder decay sequence $D(k)$, the average energy step response $S(k)$, and the impulse response maximum level envelope $L(k)$,

$$D(k) = 10\log_{10}\left(\frac{1}{N}\sum_{n=1}^{N}\sum_{l=k}^{M-1}\frac{h_n^2(l)}{\sum_{j=0}^{M-1}h_n^2(j)}\right), \quad (7)$$

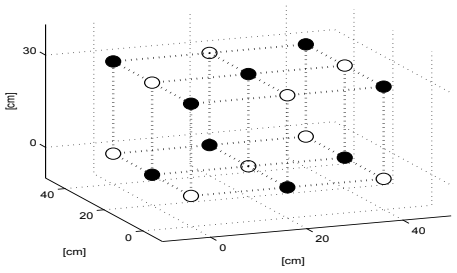

Fig. 2. Geometry of microphone positions for filter design (white) and validation (black).

$$S(k) = \frac{1}{N}\sum_{n=1}^{N}\sum_{l=0}^{k}\frac{h_n^2(l)}{\sum_{j=0}^{M-1}h_n^2(j)}, \quad (8)$$

$$L(k) = 20\log_{10}(\max_n |h_n(k)|), \quad (9)$$

defined in (7), (8) and (9) respectively for evaluation of the time domain properties. Here $h_n(k)$; $k \in \{0, ..., M-1\}$ is an impulse response of length M in microphone position n; $n \in \{1, ..., N\}$. All responses are time-aligned and normalized so that $\max|h_n(k)| = |h_n(k_0)| = 1$ for some time instant $k = k_0$. While $L(k)$ is useful as a worst case presentation of pre- or post-ringing problems, $S(k)$ and $D(k)$ indicate how good are the transient properties of the system.

4.3. Results

As the frequency response plots in Figure 3 show, the minimum phase and mixed phase equalizers exhibit identical performance in terms of magnitude response correction. This is expected, since by construction the filters differ only in phase response, see middle and bottom diagrams of Figure 3. The difference in performance between the filters is thus embedded in their phase curves, a difference which is best assessed in the time domain. Figure 4 clearly shows how the maximum peak level increases with the mixed phase design, and the maximum level of pre-ringing does not exceed the constraint that was set in the design (-60 dB before -5 msec). The energy step responses of Figure 5 and the Schroeder decay curves of Figure 6 show how the mixed phase filter substantially improves upon the pure minimum phase filter. For the full-range spectrum, the difference is most clearly visible in a short time window at the beginning of the responses. Below 300 Hz however, the improvement of mixed phase over minimum phase can be seen over a large time window.

5. CONCLUSIONS

A method for spatially robust mixed phase equalization of loudspeakers has been proposed and evaluated. By extending a minimum phase equalizer with a robust all pass inverse, substantial time-domain improvement over minimum phase equalization is obtained. It is our opinion that this result motivates a revision of the widespread conclusion that non-minimum phase RTF properties must be neglected in a robust equalizer design. Further studies of the method are needed in order to generalize the results to other acoustical environments. For example, the acoustical room parameters and the size of listener region will have an influence on the position-independence property. In the present work, mixed phase equalizer design was divided into minimum phase and all-pass inversion, where the mini-

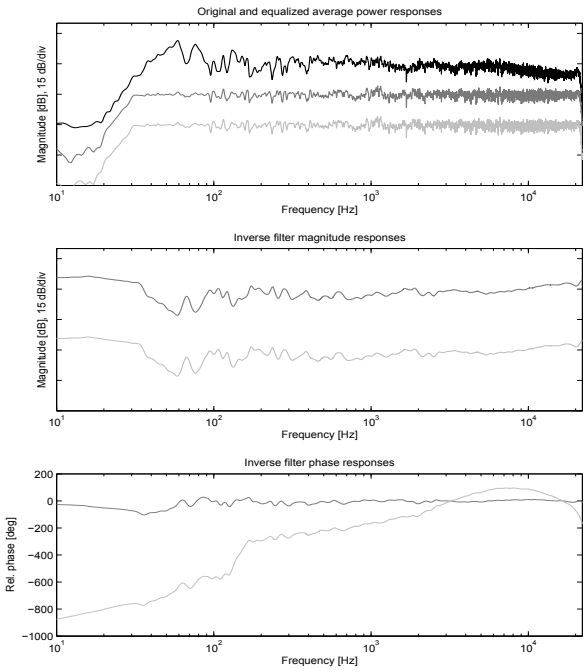

Fig. 3. Original and equalized average power responses, measured in the room (top). Magnitude and phase responses (middle and bottom, respectively) of inverse filters. Dark and light grey curves represent minimum phase and mixed phase filters, respectively.

mum phase inverse was of a rather conventional type. Further improvement is anticipated with a refined minimum phase design.

6. REFERENCES

[1] R. Wilson, "Equalization of loudspeaker drive units...," *J. Audio Eng. Soc.*, vol. 39, pp. 127–139, 1991.

[2] S.J. Elliott and P.A. Nelson, "Multiple-point equalization in a room...," *J. Audio Eng. Soc.*, vol. 37, pp. 899–907, 1989.

[3] Y. Haneda, S. Makino, and Y. Kaneda, "Multiple-point equalization of...," *IEEE Trans. on Speech and Audio Process.*, vol. 5, pp. 325–333, July 1997.

[4] P. Hatziantoniou and J. Mourjopoulos, "Generalized fractional octave...," *J. Audio Eng. Soc.*, vol. 48, pp. 259–280, April 2000.

[5] J. Mourjopoulos, "Digital equalization of room acoustics," *J. Audio Eng. Soc.*, vol. 42, pp. 884–900, November 1994.

[6] R. Greenfield and M. O. Hawksford, "Efficient filter design...," *J. Audio Eng. Soc.*, vol. 39, pp. 739–751, 1991.

[7] P. Craven and M. Gerzon, "Practical adaptive room and loudspeaker...," in *AES DSP UK Conf.* AES, 1992, pp. 121–153.

[8] M. O. Hawksford, "Digital signal processing tools for loudspeaker...," *J. Audio Eng. Soc.*, vol. 45, pp. 37–54, 1997.

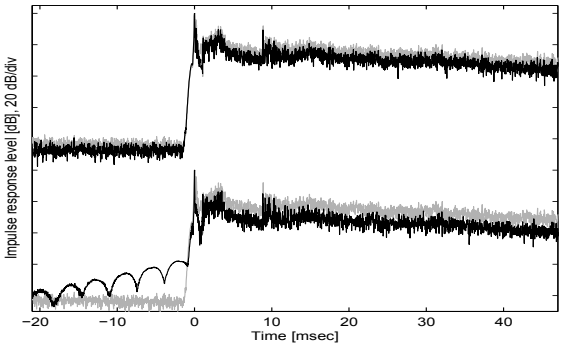

Fig. 4. Impulse response maximum level envelopes $L(k)$ of minimum phase equalization (top) and proposed method (bottom). The grey lines represent the unequalized responses.

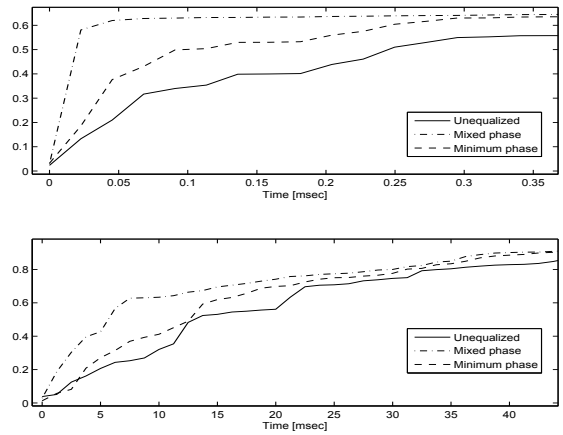

Fig. 5. Average energy step responses $S(k)$, full spectrum (top) and below 300 Hz (bottom). The curves approach unity in steady state.

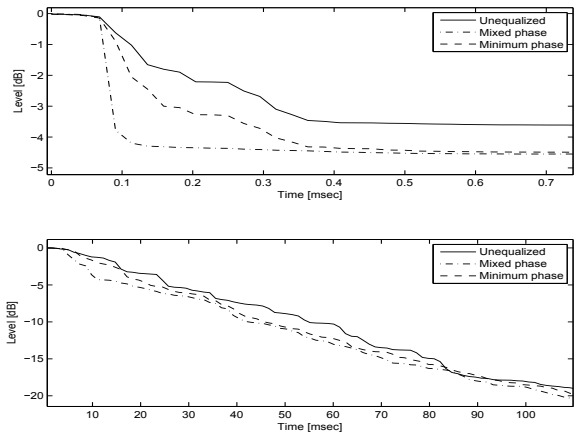

Fig. 6. Average Schroeder decay $D(k)$, full spectrum (top) and below 300 Hz (bottom).